

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: BASF Aktiengesellschaft
- (B) STREET:
- (C) CITY: Ludwigshafen
- (E) COUNTRY: Deutschland
- (F) POSTAL CODE (ZIP): 67065

(ii) TITLE OF INVENTION: Neue Poly ADP Ribose Polymerase Gene

(iii) NUMBER OF SEQUENCES: 28

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC DOS/MS DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1843 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI SENSE: NO

## (vi) ORIGINAL SOURCE:

- (F) TISSUE TYPE: Brain

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..1715
- (D) OTHER INFORMATION: /product= "Poly ADP Ribose Polymerase"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CC ATG GCG GCG CGG CGG CGA CGG AGC ACC GGC GGC GGC AGG GCG AGA	47
Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg	
1 5 10 15	
GCA TTA AAT GAA AGC AAA AGA GTT AAT AAT GGC AAC ACG GCT CCA GAA	95
Ala Leu Asn Glu Ser Lys Arg Val Asn Asn Gly Asn Thr Ala Pro Glu	
20 25 30	

GAC TCT TCC CCT GCC AAG AAA ACT CGT AGA TGC CAG AGA CAG GAG TCG Asp Ser Ser Pro Ala Lys Lys Thr Arg Arg Cys Gln Arg Gln Glu Ser	143
35 40 45	
AAA AAG ATG CCT GTG GCT GGA GGA AAA GCT AAT AAG GAC AGG ACA GAA Lys Lys Met Pro Val Ala Gly Gly Lys Ala Asn Lys Asp Arg Thr Glu	191
50 55 60	
GAC AAG CAA GAT GAA TCT GTG AAG GCC TTG CTG TTA AAG GGC AAA GCT Asp Lys Gln Asp Glu Ser Val Lys Ala Leu Leu Leu Lys Gly Lys Ala	239
65 70 75	
CCT GTG GAC CCA GAG TGT ACA GCC AAG GTG GGG AAG GCT CAT GTG TAT Pro Val Asp Pro Glu Cys Thr Ala Lys Val Gly Lys Ala His Val Tyr	287
80 85 90 95	
TGT GAA GGA AAT GAT GTC TAT GAT GTC ATG CTA AAT CAG ACC AAT CTC Cys Glu Gly Asn Asp Val Tyr Asp Val Met Leu Asn Gln Thr Asn Leu	335
100 105 110	
CAG TTC AAC AAC AAC AAG TAC TAT CTG ATT CAG CTA TTA GAA GAT GAT Gln Phe Asn Asn Asn Lys Tyr Tyr Leu Ile Gln Leu Leu Glu Asp Asp	383
115 120 125	
GCC CAG AGG AAC TTC AGT GTT TGG ATG AGA TGG GGC CGA GTT GGG AAA Ala Gln Arg Asn Phe Ser Val Trp Met Arg Trp Gly Arg Val Gly Lys	431
130 135 140	
ATG GGA CAG CAC AGC CTG GTG GCT TGT TCA GGC AAT CTC AAC AAG GCC Met Gly Gln His Ser Leu Val Ala Cys Ser Gly Asn Leu Asn Lys Ala	479
145 150 155	
AAG GAA ATC TTT CAG AAG AAA TTC CTT GAC AAA ACG AAA AAC AAT TGG Lys Glu Ile Phe Gln Lys Lys Phe Leu Asp Lys Thr Lys Asn Asn Trp	527
160 165 170 175	
GAA GAT CGA GAA AAG TTT GAG AAG GTG CCT GGA AAA TAT GAT ATG CTA Glu Asp Arg Glu Lys Phe Glu Lys Val Pro Gly Lys Tyr Asp Met Leu	575
180 185 190	
CAG ATG GAC TAT GCC ACC AAT ACT CAG GAT GAA GAG GAA ACA AAG AAA Gln Met Asp Tyr Ala Thr Asn Thr Gln Asp Glu Glu Glu Thr Lys Lys	623
195 200 205	
GAG GAA TCT CTT AAA TCT CCC TTG AAG CCA GAG TCA CAG CTA GAT CTT Glu Glu Ser Leu Lys Ser Pro Leu Lys Pro Glu Ser Gln Leu Asp Leu	671
210 215 220	
CGG GTA CAG GAG TTA ATA AAG TTG ATC TGT AAT GTT CAG GCC ATG GAA Arg Val Gln Glu Leu Ile Lys Leu Ile Cys Asn Val Gln Ala Met Glu	719
225 230 235	
GAA ATG ATG ATG GAA ATG AAG TAT AAT ACC AAG AAA GCC CCA CTT GGG Glu Met Met Met Glu Met Lys Tyr Asn Thr Lys Lys Ala Pro Leu Gly	767
240 245 250 255	
AAG CTG ACA GTG GCA CAA ATC AAG GCA GGT TAC CAG TCT CTT AAG AAG Lys Leu Thr Val Ala Gln Ile Lys Ala Gly Tyr Gln Ser Leu Lys Lys	815

260										265										270										
ATT	GAG	GAT	TGT	ATT	CGG	GCT	GGC	CAG	CAT	GGA	CGA	GCT	CTC	ATG	GAA															863
Ile	Glu	Asp	Cys	Ile	Arg	Ala	Gly	Gln	His	Gly	Arg	Ala	Leu	Met	Glu															
			275					280					285																	
GCA	TGC	AAT	GAA	TTC	TAC	ACC	AGG	ATT	CCG	CAT	GAC	TTT	GGA	CTC	CGT															911
Ala	Cys	Asn	Glu	Phe	Tyr	Thr	Arg	Ile	Pro	His	Asp	Phe	Gly	Leu	Arg															
			290				295					300																		
ACT	CCT	CCA	CTA	ATC	CGG	ACA	CAG	AAG	GAA	CTG	TCA	GAA	AAA	ATA	CAA															959
Thr	Pro	Pro	Leu	Ile	Arg	Thr	Gln	Lys	Glu	Leu	Ser	Glu	Lys	Ile	Gln															
			305				310				315																			
TTA	CTA	GAG	GCT	TTG	GGA	GAC	ATT	GAA	ATT	GCT	ATT	AAG	CTG	GTG	AAA															1007
Leu	Leu	Glu	Ala	Leu	Gly	Asp	Ile	Glu	Ile	Ala	Ile	Lys	Leu	Val	Lys															
			320			325				330					335															
ACA	GAG	CTA	CAA	AGC	CCA	GAA	CAC	CCA	TTG	GAC	CAA	CAC	TAT	AGA	AAC															1055
Thr	Glu	Leu	Gln	Ser	Pro	Glu	His	Pro	Leu	Asp	Gln	His	Tyr	Arg	Asn															
				340				345						350																
CTA	CAT	TGT	GCC	TTG	CGC	CCC	CTT	GAC	CAT	GAA	AGT	TAC	GAG	TTC	AAA															1103
Leu	His	Cys	Ala	Leu	Arg	Pro	Leu	Asp	His	Glu	Ser	Tyr	Glu	Phe	Lys															
			355					360					365																	
GTG	ATT	TCC	CAG	TAC	CTA	CAA	TCT	ACC	CAT	GCT	CCC	ACA	CAC	AGC	GAC															1151
Val	Ile	Ser	Gln	Tyr	Leu	Gln	Ser	Thr	His	Ala	Pro	Thr	His	Ser	Asp															
			370				375				380																			
TAT	ACC	ATG	ACC	TTG	CTG	GAT	TTG	TTT	GAA	GTG	GAG	AAG	GAT	GGT	GAG															1199
Tyr	Thr	Met	Thr	Leu	Leu	Asp	Leu	Phe	Glu	Val	Glu	Lys	Asp	Gly	Glu															
			385			390					395																			
AAA	GAA	GCC	TTC	AGA	GAG	GAC	CTT	CAT	AAC	AGG	ATG	CTT	CTA	TGG	CAT															1247
Lys	Glu	Ala	Phe	Arg	Glu	Asp	Leu	His	Asn	Arg	Met	Leu	Leu	Trp	His															
					405				410						415															
GGT	TCC	AGG	ATG	AGT	AAC	TGG	GTG	GGA	ATC	TTG	AGC	CAT	GGG	CTT	CGA															1295
Gly	Ser	Arg	Met	Ser	Asn	Trp	Val	Gly	Ile	Leu	Ser	His	Gly	Leu	Arg															
				420				425						430																
ATT	GCC	CCA	CCT	GAA	GCT	CCC	ATC	ACA	GGT	TAC	ATG	TTT	GGG	AAA	GGA															1343
Ile	Ala	Pro	Pro	Glu	Ala	Pro	Ile	Thr	Gly	Tyr	Met	Phe	Gly	Lys	Gly															
			435				440						445																	
ATC	TAC	TTT	GCT	GAC	ATG	TCT	TCC	AAG	AGT	GCC	AAT	TAC	TGC	TTT	GCC															1391
Ile	Tyr	Phe	Ala	Asp	Met	Ser	Ser	Lys	Ser	Ala	Asn	Tyr	Cys	Phe	Ala															
			450				455					460																		
TCT	CGC	CTA	AAG	AAT	ACA	GGA	CTG	CTG	CTC	TTA	TCA	GAG	GTA	GCT	CTA															1439
Ser	Arg	Leu	Lys	Asn	Thr	Gly	Leu	Leu	Leu	Leu	Ser	Glu	Val	Ala	Leu															
			465			470					475																			
GGT	CAG	TGT	AAT	GAA	CTA	CTA	GAG	GCC	AAT	CCT	AAG	GCC	GAA	GGA	TTG															1487
Gly	Gln	Cys	Asn	Glu	Leu	Leu	Glu	Ala	Asn	Pro	Lys	Ala	Glu	Gly	Leu															
					485				490						495															
CTT	CAA	GGT	AAA	CAT	AGC	ACC	AAG	GGG	CTG	GGC	AAG	ATG	GCT	CCC	AGT															1535

Leu	Gln	Gly	Lys	His	Ser	Thr	Lys	Gly	Leu	Gly	Lys	Met	Ala	Pro	Ser		
				500					505						510		
TCT	GCC	CAC	TTC	GTC	ACC	CTG	AAT	GGG	AGT	ACA	GTG	CCA	TTA	GGA	CCA		1583
Ser	Ala	His	Phe	Val	Thr	Leu	Asn	Gly	Ser	Thr	Val	Pro	Leu	Gly	Pro		
			515					520					525				
GCA	AGT	GAC	ACA	GGA	ATT	CTG	AAT	CCA	GAT	GGT	TAT	ACC	CTC	AAC	TAC		1631
Ala	Ser	Asp	Thr	Gly	Ile	Leu	Asn	Pro	Asp	Gly	Tyr	Thr	Leu	Asn	Tyr		
		530					535					540					
AAT	GAA	TAT	ATT	GTA	TAT	AAC	CCC	AAC	CAG	GTC	CGT	ATG	CGG	TAC	CTT		1679
Asn	Glu	Tyr	Ile	Val	Tyr	Asn	Pro	Asn	Gln	Val	Arg	Met	Arg	Tyr	Leu		
	545						550				555						
TTA	AAG	GTT	CAG	TTT	AAT	TTC	CTT	CAG	CTG	TGG	TGA	ATGTTGATAT					1725
Leu	Lys	Val	Gln	Phe	Asn	Phe	Leu	Gln	Leu	Trp	*						
	560					565				570							
TAAATAAACC	AGAGATCTGA	TCTTCAAGCA	AGAAAATAAG	CAGTGTGTGTA	CTTGTGAATT												1785
TTGTGATATT	TTATGTAATA	AAACTGTAC	AGGTCTAAAA	AAAAAAAAAA	AAAAAAAAAA												1843

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 571 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met	Ala	Ala	Arg	Arg	Arg	Arg	Ser	Thr	Gly	Gly	Gly	Arg	Ala	Arg	Ala		
1						5				10				15			
Leu	Asn	Glu	Ser	Lys	Arg	Val	Asn	Asn	Gly	Asn	Thr	Ala	Pro	Glu	Asp		
			20					25					30				
Ser	Ser	Pro	Ala	Lys	Lys	Thr	Arg	Arg	Cys	Gln	Arg	Gln	Glu	Ser	Lys		
		35					40					45					
Lys	Met	Pro	Val	Ala	Gly	Gly	Lys	Ala	Asn	Lys	Asp	Arg	Thr	Glu	Asp		
	50					55					60						
Lys	Gln	Asp	Glu	Ser	Val	Lys	Ala	Leu	Leu	Leu	Lys	Gly	Lys	Ala	Pro		
	65					70				75					80		
Val	Asp	Pro	Glu	Cys	Thr	Ala	Lys	Val	Gly	Lys	Ala	His	Val	Tyr	Cys		
				85					90					95			
Glu	Gly	Asn	Asp	Val	Tyr	Asp	Val	Met	Leu	Asn	Gln	Thr	Asn	Leu	Gln		
		100						105					110				
Phe	Asn	Asn	Asn	Lys	Tyr	Tyr	Leu	Ile	Gln	Leu	Leu	Glu	Asp	Asp	Ala		
	115						120					125					
Gln	Arg	Asn	Phe	Ser	Val	Trp	Met	Arg	Trp	Gly	Arg	Val	Gly	Lys	Met		

130					135					140					
Gly	Gln	His	Ser	Leu	Val	Ala	Cys	Ser	Gly	Asn	Leu	Asn	Lys	Ala	Lys
145					150					155					160
Glu	Ile	Phe	Gln	Lys	Lys	Phe	Leu	Asp	Lys	Thr	Lys	Asn	Asn	Trp	Glu
				165					170					175	
Asp	Arg	Glu	Lys	Phe	Glu	Lys	Val	Pro	Gly	Lys	Tyr	Asp	Met	Leu	Gln
			180					185					190		
Met	Asp	Tyr	Ala	Thr	Asn	Thr	Gln	Asp	Glu	Glu	Glu	Thr	Lys	Lys	Glu
		195					200					205			
Glu	Ser	Leu	Lys	Ser	Pro	Leu	Lys	Pro	Glu	Ser	Gln	Leu	Asp	Leu	Arg
	210					215					220				
Val	Gln	Glu	Leu	Ile	Lys	Leu	Ile	Cys	Asn	Val	Gln	Ala	Met	Glu	Glu
225					230					235					240
Met	Met	Met	Glu	Met	Lys	Tyr	Asn	Thr	Lys	Lys	Ala	Pro	Leu	Gly	Lys
				245					250					255	
Leu	Thr	Val	Ala	Gln	Ile	Lys	Ala	Gly	Tyr	Gln	Ser	Leu	Lys	Lys	Ile
			260					265					270		
Glu	Asp	Cys	Ile	Arg	Ala	Gly	Gln	His	Gly	Arg	Ala	Leu	Met	Glu	Ala
		275					280					285			
Cys	Asn	Glu	Phe	Tyr	Thr	Arg	Ile	Pro	His	Asp	Phe	Gly	Leu	Arg	Thr
	290					295					300				
Pro	Pro	Leu	Ile	Arg	Thr	Gln	Lys	Glu	Leu	Ser	Glu	Lys	Ile	Gln	Leu
305					310					315					320
Leu	Glu	Ala	Leu	Gly	Asp	Ile	Glu	Ile	Ala	Ile	Lys	Leu	Val	Lys	Thr
				325					330					335	
Glu	Leu	Gln	Ser	Pro	Glu	His	Pro	Leu	Asp	Gln	His	Tyr	Arg	Asn	Leu
			340					345					350		
His	Cys	Ala	Leu	Arg	Pro	Leu	Asp	His	Glu	Ser	Tyr	Glu	Phe	Lys	Val
		355					360					365			
Ile	Ser	Gln	Tyr	Leu	Gln	Ser	Thr	His	Ala	Pro	Thr	His	Ser	Asp	Tyr
	370					375					380				
Thr	Met	Thr	Leu	Leu	Asp	Leu	Phe	Glu	Val	Glu	Lys	Asp	Gly	Glu	Lys
385					390					395					400
Glu	Ala	Phe	Arg	Glu	Asp	Leu	His	Asn	Arg	Met	Leu	Leu	Trp	His	Gly
				405					410					415	
Ser	Arg	Met	Ser	Asn	Trp	Val	Gly	Ile	Leu	Ser	His	Gly	Leu	Arg	Ile
			420					425					430		
Ala	Pro	Pro	Glu	Ala	Pro	Ile	Thr	Gly	Tyr	Met	Phe	Gly	Lys	Gly	Ile
			435				440					445			
Tyr	Phe	Ala	Asp	Met	Ser	Ser	Lys	Ser	Ala	Asn	Tyr	Cys	Phe	Ala	Ser

450	455	460
Arg Leu Lys Asn Thr Gly Leu Leu Leu Leu Ser Glu Val Ala Leu Gly		
465	470	475 480
Gln Cys Asn Glu Leu Leu Glu Ala Asn Pro Lys Ala Glu Gly Leu Leu		
	485 490	495
Gln Gly Lys His Ser Thr Lys Gly Leu Gly Lys Met Ala Pro Ser Ser		
	500 505	510
Ala His Phe Val Thr Leu Asn Gly Ser Thr Val Pro Leu Gly Pro Ala		
	515 520	525
Ser Asp Thr Gly Ile Leu Asn Pro Asp Gly Tyr Thr Leu Asn Tyr Asn		
	530 535	540
Glu Tyr Ile Val Tyr Asn Pro Asn Gln Val Arg Met Arg Tyr Leu Leu		
	545 550	555 560
Lys Val Gln Phe Asn Phe Leu Gln Leu Trp *		
	565 570	

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI SENSE: NO

## (vi) ORIGINAL SOURCE:

- (F) TISSUE TYPE: Uterus

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 242..1843
- (D) OTHER INFORMATION: /product= "Poly ADP Ribose Polymerase"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA	60
TAGCCGATGT CTAATCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC	120
TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG	180
GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATGTCCCTGC TTTTCTTGGC	240
C ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT GAG	286
Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro Glu	

575					580					585						
AAG Lys	AAG Lys	AAG Lys	GGC Gly 590	CGG Arg	CAG Gln	GCA Ala	GGA Gly 595	AGG Arg	GAG Glu	GAG Glu	GAC Asp	CCC Pro	TTC Phe 600	CGC Arg	TCC Ser	334
ACC Thr	GCT Ala	GAG Glu 605	GCC Ala	CTC Leu	AAG Lys	GCC Ala	ATA Ile 610	CCC Pro	GCA Ala	GAG Glu	AAG Lys	CGC Arg 615	ATA Ile	ATC Ile	CGC Arg	382
GTG Val	GAT Asp 620	CCA Pro	ACA Thr	TGT Cys	CCA Pro	CTC Leu 625	AGC Ser 625	AGC Ser	AAC Asn	CCC Pro	GGG Gly 630	ACC Thr	CAG Gln	GTG Val	TAT Tyr	430
GAG Glu 635	GAC Asp	TAC Tyr	AAC Asn	TGC Cys	ACC Thr 640	CTG Leu	AAC Asn	CAG Gln	ACC Thr	AAC Asn 645	ATC Ile	GAG Glu	AAC Asn	AAC Asn	AAC Asn 650	478
AAC Asn	AAG Lys	TTC Phe	TAC Tyr	ATC Ile 655	ATC Ile	CAG Gln	CTG Leu	CTC Leu	CAA Gln 660	GAC Asp	AGC Ser	AAC Asn	CGC Arg	TTC Phe 665	TTC Phe	526
ACC Thr	TGC Cys	TGG Trp 670	AAC Asn	CGC Arg	TGG Trp	GGC Gly	CGT Arg 675	GTG Val 675	GGA Gly	GAG Glu	GTC Val	GGC Gly 680	CAG Gln	TCA Ser	AAG Lys	574
ATC Ile	AAC Asn 685	CAC His	TTC Phe	ACA Thr	AGG Arg	CTA Leu 690	GAA Glu 690	GAT Asp	GCA Ala	AAG Lys	AAG Lys	GAC Asp 695	TTT Phe	GAG Glu	AAG Lys	622
AAA Lys 700	TTT Phe	CGG Arg	GAA Glu	AAG Lys	ACC Thr 705	AAG Lys 705	AAC Asn	AAC Asn	TGG Trp	GCA Ala	GAG Glu 710	CGG Arg	GAC Asp	CAC His	TTT Phe	670
GTG Val 715	TCT Ser	CAC His	CCG Pro	GGC Gly 720	AAG Lys 720	TAC Tyr	ACA Thr	CTT Leu	ATC Ile	GAA Glu 725	GTA Val	CAG Gln	GCA Ala	GAG Glu	GAT Asp 730	718
GAG Glu	GCC Ala	CAG Gln	GAA Glu 735	GCT Ala	GTG Val 735	GTG Val	AAG Lys	GTG Val	GAC Asp 740	AGA Arg	GGC Gly	CCA Pro	GTG Val	AGG Arg	ACT Thr	766
GTG Val	ACT Thr	AAG Lys 750	CGG Arg	GTG Val	CAG Gln	CCC Pro	TGC Cys 755	TCC Ser 755	CTG Leu	GAC Asp	CCA Pro	GCC Ala	ACG Thr 760	CAG Gln	AAG Lys	814
CTC Leu	ATC Ile	ACT Thr 765	AAC Asn	ATC Ile	TTC Phe	AGC Ser	AAG Lys 770	GAG Glu 770	ATG Met	TTC Phe	AAG Lys	AAC Asn 775	ACC Thr	ATG Met	GCC Ala	862
CTC Leu	ATG Met 780	GAC Asp	CTG Leu	GAT Asp	GTG Val	AAG Lys 785	AAG Lys	ATG Met	CCC Pro	CTG Leu	GGA Gly 790	AAG Lys	CTG Leu	AGC Ser	AAG Lys	910
CAA Gln 795	CAG Gln	ATT Ile	GCA Ala	CGG Arg	GGT Gly 800	TTC Phe	GAG Glu	GCC Ala	TTG Leu	GAG Glu 805	GCG Ala	CTG Leu	GAG Glu	GAG Glu	GCC Ala 810	958
CTG	AAA	GGC	CCC	ACG	GAT	GGT	GGC	CAA	AGC	CTG	GAG	GAG	CTG	TCC	TCA	1006

Leu	Lys	Gly	Pro	Thr	Asp	Gly	Gly	Gln	Ser	Leu	Glu	Glu	Leu	Ser	Ser		
				815					820					825			
CAC	TTT	TAC	ACC	GTC	ATC	CCG	CAC	AAC	TTC	GGC	CAC	AGC	CAG	CCC	CCG		1054
His	Phe	Tyr	Thr	Val	Ile	Pro	His	Asn	Phe	Gly	His	Ser	Gln	Pro	Pro		
			830					835					840				
CCC	ATC	AAT	TCC	CCT	GAG	CTT	CTG	CAG	GCC	AAG	AAG	GAC	ATG	CTG	CTG		1102
Pro	Ile	Asn	Ser	Pro	Glu	Leu	Leu	Gln	Ala	Lys	Lys	Asp	Met	Leu	Leu		
		845					850					855					
GTG	CTG	GCG	GAC	ATC	GAG	CTG	GCC	CAG	GCC	CTG	CAG	GCA	GTC	TCT	GAG		1150
Val	Leu	Ala	Asp	Ile	Glu	Leu	Ala	Gln	Ala	Leu	Gln	Ala	Val	Ser	Glu		
	860					865					870						
CAG	GAG	AAG	ACG	GTG	GAG	GAG	GTG	CCA	CAC	CCC	CTG	GAC	CGA	GAC	TAC		1198
Gln	Glu	Lys	Thr	Val	Glu	Glu	Val	Pro	His	Pro	Leu	Asp	Arg	Asp	Tyr		
875				880				885							890		
CAG	CTT	CTC	AAG	TGC	CAG	CTG	CAG	CTG	CTA	GAC	TCT	GGA	GCA	CCT	GAG		1246
Gln	Leu	Leu	Lys	Cys	Gln	Leu	Gln	Leu	Leu	Asp	Ser	Gly	Ala	Pro	Glu		
				895				900						905			
TAC	AAG	GTG	ATA	CAG	ACC	TAC	TTA	GAA	CAG	ACT	GGC	AGC	AAC	CAC	AGG		1294
Tyr	Lys	Val	Ile	Gln	Thr	Tyr	Leu	Glu	Gln	Thr	Gly	Ser	Asn	His	Arg		
			910					915					920				
TGC	CCT	ACA	CTT	CAA	CAC	ATC	TGG	AAA	GTA	AAC	CAA	GAA	GGG	GAG	GAA		1342
Cys	Pro	Thr	Leu	Gln	His	Ile	Trp	Lys	Val	Asn	Gln	Glu	Gly	Glu	Glu		
		925					930					935					
GAC	AGA	TTC	CAG	GCC	CAC	TCC	AAA	CTG	GGT	AAT	CGG	AAG	CTG	CTG	TGG		1390
Asp	Arg	Phe	Gln	Ala	His	Ser	Lys	Leu	Gly	Asn	Arg	Lys	Leu	Leu	Trp		
	940					945					950						
CAT	GGC	ACC	AAC	ATG	GCC	GTG	GTG	GCC	GCC	ATC	CTC	ACT	AGT	GGG	CTC		1438
His	Gly	Thr	Asn	Met	Ala	Val	Val	Ala	Ala	Ile	Leu	Thr	Ser	Gly	Leu		
955					960					965					970		
CGC	ATC	ATG	CCA	CAT	TCT	GGT	GGG	CGT	GTT	GGC	AAG	GGC	ATC	TAC	TTT		1486
Arg	Ile	Met	Pro	His	Ser	Gly	Gly	Arg	Val	Gly	Lys	Gly	Ile	Tyr	Phe		
				975				980						985			
GCC	TCA	GAG	AAC	AGC	AAG	TCA	GCT	GGA	TAT	GTT	ATT	GGC	ATG	AAG	TGT		1534
Ala	Ser	Glu	Asn	Ser	Lys	Ser	Ala	Gly	Tyr	Val	Ile	Gly	Met	Lys	Cys		
			990					995					1000				
GGG	GCC	CAC	CAT	GTC	GGC	TAC	ATG	TTC	CTG	GGT	GAG	GTG	GCC	CTG	GGC		1582
Gly	Ala	His	His	Val	Gly	Tyr	Met	Phe	Leu	Gly	Glu	Val	Ala	Leu	Gly		
		1005					1010					1015					
AGA	GAG	CAC	CAT	ATC	AAC	ACG	GAC	AAC	CCC	AGC	TTG	AAG	AGC	CCA	CCT		1630
Arg	Glu	His	His	Ile	Asn	Thr	Asp	Asn	Pro	Ser	Leu	Lys	Ser	Pro	Pro		
	1020					1025					1030						
CCT	GGC	TTC	GAC	AGT	GTC	ATT	GCC	CGA	GGC	CAC	ACC	GAG	CCT	GAT	CCG		1678
Pro	Gly	Phe	Asp	Ser	Val	Ile	Ala	Arg	Gly	His	Thr	Glu	Pro	Asp	Pro		
1035					1040					1045					1050		



ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG CCC	1726
Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro	
1055 1060 1065	
CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC TCC	1774
Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser	
1070 1075 1080	
CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC TAC	1822
Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr	
1085 1090 1095	
CTG CTG GAG GTC CAC CTC TGA GTGCCCCGCC TGTCCCCCGG GGTCTGCAA	1873
Leu Leu Glu Val His Leu *	
1100 1105	
GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT	1933
TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA	1993
CTTATGCCTC CTAAGTAAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC	2053
CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC	2113
ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT	2173
AATGTGCACC GGGTTCACCC ATGTTTTTCAT AAATGACAAG ATTTCTCCT TAAAAA	2233
AAAAA	2265

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 534 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Ala	Pro	Lys	Pro	Lys	Pro	Trp	Val	Gln	Thr	Glu	Gly	Pro	Glu	Lys
1				5					10					15	
Lys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu	Glu	Asp	Pro	Phe	Arg	Ser	Thr
		20						25					30		
Ala	Glu	Ala	Leu	Lys	Ala	Ile	Pro	Ala	Glu	Lys	Arg	Ile	Ile	Arg	Val
		35					40						45		
Asp	Pro	Thr	Cys	Pro	Leu	Ser	Asn	Pro	Gly	Thr	Gln	Val	Tyr	Glu	
		50				55				60					
Asp	Tyr	Asn	Cys	Thr	Leu	Asn	Gln	Thr	Asn	Ile	Glu	Asn	Asn	Asn	Asn
65					70				75					80	
Lys	Phe	Tyr	Ile	Ile	Gln	Leu	Leu	Gln	Asp	Ser	Asn	Arg	Phe	Phe	Thr
			85					90						95	
Cys	Trp	Asn	Arg	Trp	Gly	Arg	Val	Gly	Glu	Val	Gly	Gln	Ser	Lys	Ile

100					105					110						
Asn	His	Phe	Thr	Arg	Leu	Glu	Asp	Ala	Lys	Lys	Asp	Phe	Glu	Lys	Lys	
115					120					125						
Phe	Arg	Glu	Lys	Thr	Lys	Asn	Asn	Trp	Ala	Glu	Arg	Asp	His	Phe	Val	
130					135					140						
Ser	His	Pro	Gly	Lys	Tyr	Thr	Leu	Ile	Glu	Val	Gln	Ala	Glu	Asp	Glu	
145					150					155					160	
Ala	Gln	Glu	Ala	Val	Val	Lys	Val	Asp	Arg	Gly	Pro	Val	Arg	Thr	Val	
165					170					175						
Thr	Lys	Arg	Val	Gln	Pro	Cys	Ser	Leu	Asp	Pro	Ala	Thr	Gln	Lys	Leu	
180					185					190						
Ile	Thr	Asn	Ile	Phe	Ser	Lys	Glu	Met	Phe	Lys	Asn	Thr	Met	Ala	Leu	
195					200					205						
Met	Asp	Leu	Asp	Val	Lys	Lys	Met	Pro	Leu	Gly	Lys	Leu	Ser	Lys	Gln	
210					215					220						
Gln	Ile	Ala	Arg	Gly	Phe	Glu	Ala	Leu	Glu	Ala	Leu	Glu	Glu	Ala	Leu	
225					230					235					240	
Lys	Gly	Pro	Thr	Asp	Gly	Gly	Gln	Ser	Leu	Glu	Glu	Leu	Ser	Ser	His	
245					250					255						
Phe	Tyr	Thr	Val	Ile	Pro	His	Asn	Phe	Gly	His	Ser	Gln	Pro	Pro	Pro	
260					265					270						
Ile	Asn	Ser	Pro	Glu	Leu	Leu	Gln	Ala	Lys	Lys	Asp	Met	Leu	Leu	Val	
275					280					285						
Leu	Ala	Asp	Ile	Glu	Leu	Ala	Gln	Ala	Leu	Gln	Ala	Val	Ser	Glu	Gln	
290					295					300						
Glu	Lys	Thr	Val	Glu	Glu	Val	Pro	His	Pro	Leu	Asp	Arg	Asp	Tyr	Gln	
305					310					315					320	
Leu	Leu	Lys	Cys	Gln	Leu	Gln	Leu	Leu	Asp	Ser	Gly	Ala	Pro	Glu	Tyr	
325					330					335						
Lys	Val	Ile	Gln	Thr	Tyr	Leu	Glu	Gln	Thr	Gly	Ser	Asn	His	Arg	Cys	
340					345					350						
Pro	Thr	Leu	Gln	His	Ile	Trp	Lys	Val	Asn	Gln	Glu	Gly	Glu	Glu	Asp	
355					360					365						
Arg	Phe	Gln	Ala	His	Ser	Lys	Leu	Gly	Asn	Arg	Lys	Leu	Leu	Trp	His	
370					375					380						
Gly	Thr	Asn	Met	Ala	Val	Val	Ala	Ala	Ile	Leu	Thr	Ser	Gly	Leu	Arg	
385					390					395					400	
Ile	Met	Pro	His	Ser	Gly	Gly	Arg	Val	Gly	Lys	Gly	Ile	Tyr	Phe	Ala	
405					410					415						
Ser	Glu	Asn	Ser	Lys	Ser	Ala	Gly	Tyr	Val	Ile	Gly	Met	Lys	Cys	Gly	

420	425	430
Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Arg		
435	440	445
Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro Pro		
450	455	460
Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro Thr		
465	470	475
Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro Gln		
485	490	495
Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser Gln		
500	505	510
Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr Leu		
515	520	525
Leu Glu Val His Leu *		
530		

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI SENSE: NO

## (vi) ORIGINAL SOURCE:

(F) TISSUE TYPE: Uterus

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 221..1843
- (D) OTHER INFORMATION: /product= "Poly ADP Ribose Polymerase"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA	60
TAGCCGATGT CTAATCCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC	120
TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG	180
GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATG TCC CTG CTT TTC	235
Met Ser Leu Leu Phe	
535	
TTG GCC ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT	283

Leu	Ala	Met	Ala	Pro	Lys	Pro	Lys	Pro	Trp	Val	Gln	Thr	Glu	Gly	Pro		
540					545					550					555		
GAG	AAG	AAG	AAG	GGC	CGG	CAG	GCA	GGA	AGG	GAG	GAG	GAC	CCC	TTC	CGC	331	
Glu	Lys	Lys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu	Glu	Asp	Pro	Phe	Arg		
				560					565					570			
TCC	ACC	GCT	GAG	GCC	CTC	AAG	GCC	ATA	CCC	GCA	GAG	AAG	CGC	ATA	ATC	379	
Ser	Thr	Ala	Glu	Ala	Leu	Lys	Ala	Ile	Pro	Ala	Glu	Lys	Arg	Ile	Ile		
			575					580					585				
CGC	GTG	GAT	CCA	ACA	TGT	CCA	CTC	AGC	AGC	AAC	CCC	GGG	ACC	CAG	GTG	427	
Arg	Val	Asp	Pro	Thr	Cys	Pro	Leu	Ser	Ser	Asn	Pro	Gly	Thr	Gln	Val		
		590					595					600					
TAT	GAG	GAC	TAC	AAC	TGC	ACC	CTG	AAC	CAG	ACC	AAC	ATC	GAG	AAC	AAC	475	
Tyr	Glu	Asp	Tyr	Asn	Cys	Thr	Leu	Asn	Gln	Thr	Asn	Ile	Glu	Asn	Asn		
	605					610					615						
AAC	AAC	AAG	TTC	TAC	ATC	ATC	CAG	CTG	CTC	CAA	GAC	AGC	AAC	CGC	TTC	523	
Asn	Asn	Lys	Phe	Tyr	Ile	Ile	Gln	Leu	Leu	Gln	Asp	Ser	Asn	Arg	Phe		
	620				625					630					635		
TTC	ACC	TGC	TGG	AAC	CGC	TGG	GGC	CGT	GTG	GGA	GAG	GTC	GGC	CAG	TCA	571	
Phe	Thr	Cys	Trp	Asn	Arg	Trp	Gly	Arg	Val	Gly	Glu	Val	Gly	Gln	Ser		
				640				645						650			
AAG	ATC	AAC	CAC	TTC	ACA	AGG	CTA	GAA	GAT	GCA	AAG	AAG	GAC	TTT	GAG	619	
Lys	Ile	Asn	His	Phe	Thr	Arg	Leu	Glu	Asp	Ala	Lys	Lys	Asp	Phe	Glu		
			655				660						665				
AAG	AAA	TTT	CGG	GAA	AAG	ACC	AAG	AAC	AAC	TGG	GCA	GAG	CGG	GAC	CAC	667	
Lys	Lys	Phe	Arg	Glu	Lys	Thr	Lys	Asn	Asn	Trp	Ala	Glu	Arg	Asp	His		
		670					675					680					
TTT	GTG	TCT	CAC	CCG	GGC	AAG	TAC	ACA	CTT	ATC	GAA	GTA	CAG	GCA	GAG	715	
Phe	Val	Ser	His	Pro	Gly	Lys	Tyr	Thr	Leu	Ile	Glu	Val	Gln	Ala	Glu		
	685					690					695						
GAT	GAG	GCC	CAG	GAA	GCT	GTG	GTG	AAG	GTG	GAC	AGA	GGC	CCA	GTG	AGG	763	
Asp	Glu	Ala	Gln	Glu	Ala	Val	Val	Lys	Val	Asp	Arg	Gly	Pro	Val	Arg		
	700				705					710					715		
ACT	GTG	ACT	AAG	CGG	GTG	CAG	CCC	TGC	TCC	CTG	GAC	CCA	GCC	ACG	CAG	811	
Thr	Val	Thr	Lys	Arg	Val	Gln	Pro	Cys	Ser	Leu	Asp	Pro	Ala	Thr	Gln		
				720				725						730			
AAG	CTC	ATC	ACT	AAC	ATC	TTC	AGC	AAG	GAG	ATG	TTC	AAG	AAC	ACC	ATG	859	
Lys	Leu	Ile	Thr	Asn	Ile	Phe	Ser	Lys	Glu	Met	Phe	Lys	Asn	Thr	Met		
			735				740						745				
GCC	CTC	ATG	GAC	CTG	GAT	GTG	AAG	AAG	ATG	CCC	CTG	GGA	AAG	CTG	AGC	907	
Ala	Leu	Met	Asp	Leu	Asp	Val	Lys	Lys	Met	Pro	Leu	Gly	Lys	Leu	Ser		
		750					755					760					
AAG	CAA	CAG	ATT	GCA	CGG	GGT	TTC	GAG	GCC	TTG	GAG	GCG	CTG	GAG	GAG	955	
Lys	Gln	Gln	Ile	Ala	Arg	Gly	Phe	Glu	Ala	Leu	Glu	Ala	Leu	Glu	Glu		
	765					770					775						

GCC CTG AAA GGC CCC ACG GAT GGT GGC CAA AGC CTG GAG GAG CTG TCC	1003
Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser	
780 785 790 795	
TCA CAC TTT TAC ACC GTC ATC CCG CAC AAC TTC GGC CAC AGC CAG CCC	1051
Ser His Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro	
800 805 810	
CCG CCC ATC AAT TCC CCT GAG CTT CTG CAG GCC AAG AAG GAC ATG CTG	1099
Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu	
815 820 825	
CTG GTG CTG GCG GAC ATC GAG CTG GCC CAG GCC CTG CAG GCA GTC TCT	1147
Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser	
830 835 840	
GAG CAG GAG AAG ACG GTG GAG GAG GTG CCA CAC CCC CTG GAC CGA GAC	1195
Glu Gln Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp	
845 850 855	
TAC CAG CTT CTC AAG TGC CAG CTG CAG CTG CTA GAC TCT GGA GCA CCT	1243
Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro	
860 865 870 875	
GAG TAC AAG GTG ATA CAG ACC TAC TTA GAA CAG ACT GGC AGC AAC CAC	1291
Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His	
880 885 890	
AGG TGC CCT ACA CTT CAA CAC ATC TGG AAA GTA AAC CAA GAA GGG GAG	1339
Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu	
895 900 905	
GAA GAC AGA TTC CAG GCC CAC TCC AAA CTG GGT AAT CGG AAG CTG CTG	1387
Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu	
910 915 920	
TGG CAT GGC ACC AAC ATG GCC GTG GTG GCC GCC ATC CTC ACT AGT GGG	1435
Trp His Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly	
925 930 935	
CTC CGC ATC ATG CCA CAT TCT GGT GGG CGT GTT GGC AAG GGC ATC TAC	1483
Leu Arg Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr	
940 945 950 955	
TTT GCC TCA GAG AAC AGC AAG TCA GCT GGA TAT GTT ATT GGC ATG AAG	1531
Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys	
960 965 970	
TGT GGG GCC CAC CAT GTC GGC TAC ATG TTC CTG GGT GAG GTG GCC CTG	1579
Cys Gly Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu	
975 980 985	
GGC AGA GAG CAC CAT ATC AAC ACG GAC AAC CCC AGC TTG AAG AGC CCA	1627
Gly Arg Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro	
990 995 1000	
CCT CCT GGC TTC GAC AGT GTC ATT GCC CGA GGC CAC ACC GAG CCT GAT	1675
Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp	

1005	1010	1015	
CCG ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG			1723
Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val			
1020	1025	1030	1035
CCC CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC			1771
Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe			
	1040	1045	1050
TCC CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC			1819
Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg			
	1055	1060	1065
TAC CTG CTG GAG GTC CAC CTC TGA GTGCCCCGCC TGTCCCCCGG GGTCTGCAA			1873
Tyr Leu Leu Glu Val His Leu *			
	1070	1075	
GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT			1933
TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA			1993
CTTATGCCTC CTAAGTAAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC			2053
CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC			2113
ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT			2173
AATGTGCACC GGGTTCACCC ATGTTTTTCAT AAATGACAAG ATTTCTCCTT TAAAAA			2233
AAAAA			2265

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 541 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met	Ser	Leu	Leu	Phe	Leu	Ala	Met	Ala	Pro	Lys	Pro	Lys	Pro	Trp	Val
1					5				10					15	
Gln	Thr	Glu	Gly	Pro	Glu	Lys	Lys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu
		20						25					30		
Glu	Asp	Pro	Phe	Arg	Ser	Thr	Ala	Glu	Ala	Leu	Lys	Ala	Ile	Pro	Ala
		35					40					45			
Glu	Lys	Arg	Ile	Ile	Arg	Val	Asp	Pro	Thr	Cys	Pro	Leu	Ser	Ser	Asn
	50					55					60				
Pro	Gly	Thr	Gln	Val	Tyr	Glu	Asp	Tyr	Asn	Cys	Thr	Leu	Asn	Gln	Thr
65					70					75				80	
Asn	Ile	Glu	Asn	Asn	Asn	Asn	Lys	Phe	Tyr	Ile	Ile	Gln	Leu	Leu	Gln

- 62 -

405	410	415
Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr		
420	425	430
Val Ile Gly Met Lys Cys Gly Ala His His Val Gly Tyr Met Phe Leu		
435	440	445
Gly Glu Val Ala Leu Gly Arg Glu His His Ile Asn Thr Asp Asn Pro		
450	455	460
Ser Leu Lys Ser Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly		
465	470	475
His Thr Glu Pro Asp Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly		
485	490	495
Gln Gln Val Val Val Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe		
500	505	510
Ser Ser Ser Thr Phe Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser		
515	520	525
Gln Cys Arg Leu Arg Tyr Leu Leu Glu Val His Leu *		
530	535	540

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1740 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mus musculus

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 112..1710

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCCGGCTTTC ACTTTTCTG CTGCCTCGGG GAACACCTCG AGCCAACTGC TTCCTAACTC	60
AGGGTGGGCA GAACTGACGG GATCTAAGCT TCTGCATCTC TGAGGAGAAC C ATG GCT	117
Met Ala	
CCA AAA CGA AAG GCC TCT GTG CAG ACT GAG GGC TCC AAG AAG CAG CGA	165
Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys Gln Arg	
545                      550                      555	



CAA GGG ACA GAG GAG GAG GAC AGC TTC CGG TCC ACT GCC GAG GCT CTC	213
Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu Ala Leu	
560 565 570 575	
AGA GCA GCA CCT GCT GAT AAT CGG GTC ATC CGT GTG GAC CCC TCA TGT	261
Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro Ser Cys	
580 585 590	
CCA TTC AGC CGG AAC CCC GGG ATA CAG GTC CAC GAG GAC TAT GAC TGT	309
Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr Asp Cys	
595 600 605	
ACC CTG AAC CAG ACC AAC ATC GGC AAC AAC AAC AAC AAG TTC TAT ATT	357
Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe Tyr Ile	
610 615 620	
ATC CAA CTG CTG GAG GAG GGT AGT CGC TTC TTC TGC TGG AAT CGC TGG	405
Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn Arg Trp	
625 630 635	
GGC CGC GTG GGA GAG GTG GGC CAG AGC AAG ATG AAC CAC TTC ACC TGC	453
Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe Thr Cys	
640 645 650 655	
CTG GAA GAT GCA AAG AAG GAC TTT AAG AAG AAA TTT TGG GAG AAG ACT	501
Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu Lys Thr	
660 665 670	
AAA AAC AAA TGG GAG GAG CGG GAC CGT TTT GTG GCC CAG CCC AAC AAG	549
Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro Asn Lys	
675 680 685	
TAC ACA CTT ATA GAA GTC CAG GGA GAA GCA GAG AGC CAA GAG GCT GTA	597
Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu Ala Val	
690 695 700	
GTG AAG GCC TTA TCT CCC CAG GTG GAC AGC GGC CCT GTG AGG ACC GTG	645
Val Lys Ala Leu Ser Pro Gln Val Asp Ser Gly Pro Val Arg Thr Val	
705 710 715	
GTC AAG CCC TGC TCC CTA GAC CCT GCC ACC CAG AAC CTT ATC ACC AAC	693
Val Lys Pro Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn	
720 725 730 735	
ATC TTC AGC AAA GAG ATG TTC AAG AAC GCA ATG ACC CTC ATG AAC CTG	741
Ile Phe Ser Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu	
740 745 750	
GAT GTG AAG AAG ATG CCC TTG GGA AAG CTG ACC AAG CAG CAG ATT GCC	789
Asp Val Lys Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala	
755 760 765	
CGT GGC TTC GAG GCC TTG GAA GCT CTA GAG GAG GCC ATG AAA AAC CCC	837
Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro	
770 775 780	
ACA GGG GAT GGC CAG AGC CTG GAA GAG CTC TCC TCC TGC TTC TAC ACT	885
Thr Gly Asp Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr	

785					790					795						
GTC	ATC	CCA	CAC	AAC	TTC	GGC	CGC	AGC	CGA	CCC	CCG	CCC	ATC	AAC	TCC	933
Val	Ile	Pro	His	Asn	Phe	Gly	Arg	Ser	Arg	Pro	Pro	Pro	Ile	Asn	Ser	
800					805					810					815	
CCT	GAT	GTG	CTT	CAG	GCC	AAG	AAG	GAC	ATG	CTG	CTG	GTG	CTA	GCG	GAC	981
Pro	Asp	Val	Leu	Gln	Ala	Lys	Lys	Asp	Met	Leu	Leu	Val	Leu	Ala	Asp	
				820					825					830		
ATC	GAG	TTG	GCG	CAG	ACC	TTG	CAG	GCA	GCC	CCT	GGG	GAG	GAG	GAG	GAG	1029
Ile	Glu	Leu	Ala	Gln	Thr	Leu	Gln	Ala	Ala	Pro	Gly	Glu	Glu	Glu	Glu	
			835					840					845			
AAA	GTG	GAA	GAG	GTG	CCA	CAC	CCA	CTG	GAT	CGA	GAC	TAC	CAG	CTC	CTC	1077
Lys	Val	Glu	Glu	Val	Pro	His	Pro	Leu	Asp	Arg	Asp	Tyr	Gln	Leu	Leu	
		850					855					860				
AGG	TGC	CAG	CTT	CAA	CTG	CTG	GAC	TCC	GGG	GAG	TCC	GAG	TAC	AAG	GCA	1125
Arg	Cys	Gln	Leu	Gln	Leu	Leu	Asp	Ser	Gly	Glu	Ser	Glu	Tyr	Lys	Ala	
	865					870					875					
ATA	CAG	ACC	TAC	CTG	AAA	CAG	ACT	GGC	AAC	AGC	TAC	AGG	TGC	CCA	AAC	1173
Ile	Gln	Thr	Tyr	Leu	Lys	Gln	Thr	Gly	Asn	Ser	Tyr	Arg	Cys	Pro	Asn	
880					885					890					895	
CTG	CGG	CAT	GTT	TGG	AAA	GTG	AAC	CGA	GAA	GGG	GAG	GGA	GAC	AGG	TTC	1221
Leu	Arg	His	Val	Trp	Lys	Val	Asn	Arg	Glu	Gly	Glu	Gly	Asp	Arg	Phe	
				900					905					910		
CAG	GCC	CAC	TCC	AAA	CTG	GGC	AAT	CGG	AGG	CTG	CTG	TGG	CAC	GGC	ACC	1269
Gln	Ala	His	Ser	Lys	Leu	Gly	Asn	Arg	Arg	Leu	Leu	Trp	His	Gly	Thr	
			915					920					925			
AAT	GTG	GCC	GTG	GTG	GCT	GCC	ATC	CTC	ACC	AGT	GGG	CTC	CGA	ATC	ATG	1317
Asn	Val	Ala	Val	Val	Ala	Ala	Ile	Leu	Thr	Ser	Gly	Leu	Arg	Ile	Met	
		930					935					940				
CCA	CAC	TCG	GGT	GGT	CGT	GTT	GGC	AAG	GGT	ATT	TAT	TTT	GCC	TCT	GAG	1365
Pro	His	Ser	Gly	Gly	Arg	Val	Gly	Lys	Gly	Ile	Tyr	Phe	Ala	Ser	Glu	
	945					950					955					
AAC	AGC	AAG	TCA	GCT	GGC	TAT	GTT	ACC	ACC	ATG	CAC	TGT	GGG	GGC	CAC	1413
Asn	Ser	Lys	Ser	Ala	Gly	Tyr	Val	Thr	Thr	Met	His	Cys	Gly	Gly	His	
960					965					970					975	
CAG	GTG	GGC	TAC	ATG	TTC	CTG	GGC	GAG	GTG	GCC	CTC	GGC	AAA	GAG	CAC	1461
Gln	Val	Gly	Tyr	Met	Phe	Leu	Gly	Glu	Val	Ala	Leu	Gly	Lys	Glu	His	
				980					985					990		
CAC	ATC	ACC	ATC	GAT	GAC	CCC	AGC	TTG	AAG	AGT	CCA	CCC	CCT	GGC	TTT	1509
His	Ile	Thr	Ile	Asp	Asp	Pro	Ser	Leu	Lys	Ser	Pro	Pro	Pro	Gly	Phe	
			995					1000					1005			
GAC	AGC	GTC	ATC	GCC	CGA	GGC	CAA	ACC	GAG	CCG	GAT	CCC	GCC	CAG	GAC	1557
Asp	Ser	Val	Ile	Ala	Arg	Gly	Gln	Thr	Glu	Pro	Asp	Pro	Ala	Gln	Asp	
		1010					1015					1020				
ATT	GAA	CTT	GAA	CTG	GAT	GGG	CAG	CCG	GTG	GTG	GTG	CCC	CAA	GGC	CCG	1605

Ile Glu Leu Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln Gly Pro  
 1025 1030 1035  
 CCT GTG CAG TGC CCG TCA TTC AAA AGC TCC AGC TTC AGC CAG AGT GAA 1653  
 Pro Val Gln Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln Ser Glu  
 1040 1045 1050 1055  
 TAC CTC ATA TAC AAG GAG AGC CAG TGT CGC CTG CGC TAC CTG CTG GAG 1701  
 Tyr Leu Ile Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu Leu Glu  
 1060 1065 1070  
 ATT CAC CTC TAAGCTGCTT GCCCTCCCTA GGTCCAAGCC 1740  
 Ile His Leu

## (2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 533 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys  
 1 5 10 15  
 Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu  
 20 25 30  
 Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro  
 35 40 45  
 Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr  
 50 55 60  
 Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe  
 65 70 75 80  
 Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn  
 85 90 95  
 Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe  
 100 105 110  
 Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu  
 115 120 125  
 Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro  
 130 135 140  
 Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu  
 145 150 155 160  
 Ala Val Val Lys Ala Leu Ser Pro Gln Val Asp Ser Gly Pro Val Arg  
 165 170 175  
 Thr Val Val Lys Pro Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile

- 67 -

500

505

510

Ser Glu Tyr Leu Ile Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu  
 515 520 525

Leu Glu Ile His Leu  
 530

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1587 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mus musculus

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1584

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ATG GCT CCA AAA CGA AAG GCC TCT GTG CAG ACT GAG GGC TCC AAG AAG	48
Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys	
535 540 545	
CAG CGA CAA GGG ACA GAG GAG GAG GAC AGC TTC CGG TCC ACT GCC GAG	96
Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu	
550 555 560 565	
GCT CTC AGA GCA GCA CCT GCT GAT AAT CGG GTC ATC CGT GTG GAC CCC	144
Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro	
570 575 580	
TCA TGT CCA TTC AGC CGG AAC CCC GGG ATA CAG GTC CAC GAG GAC TAT	192
Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr	
585 590 595	
GAC TGT ACC CTG AAC CAG ACC AAC ATC GGC AAC AAC AAC AAC AAG TTC	240
Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe	
600 605 610	
TAT ATT ATC CAA CTG CTG GAG GAG GGT AGT CGC TTC TTC TGC TGG AAT	288
Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn	
615 620 625	
CGC TGG GGC CGC GTG GGA GAG GTG GGC CAG AGC AAG ATG AAC CAC TTC	336
Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe	
630 635 640 645	

ACC TGC CTG GAA GAT GCA AAG AAG GAC TTT AAG AAG AAA TTT TGG GAG	384
Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu	
650 655 660	
AAG ACT AAA AAC AAA TGG GAG GAG CGG GAC CGT TTT GTG GCC CAG CCC	432
Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro	
665 670 675	
AAC AAG TAC ACA CTT ATA GAA GTC CAG GGA GAA GCA GAG AGC CAA GAG	480
Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu	
680 685 690	
GCT GTA GTG AAG GTG GAC AGC GGC CCT GTG AGG ACC GTG GTC AAG CCC	528
Ala Val Val Lys Val Asp Ser Gly Pro Val Arg Thr Val Val Lys Pro	
695 700 705	
TGC TCC CTA GAC CCT GCC ACC CAG AAC CTT ATC ACC AAC ATC TTC AGC	576
Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn Ile Phe Ser	
710 715 720 725	
AAA GAG ATG TTC AAG AAC GCA ATG ACC CTC ATG AAC CTG GAT GTG AAG	624
Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu Asp Val Lys	
730 735 740	
AAG ATG CCC TTG GGA AAG CTG ACC AAG CAG CAG ATT GCC CGT GGC TTC	672
Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala Arg Gly Phe	
745 750 755	
GAG GCC TTG GAA GCT CTA GAG GAG GCC ATG AAA AAC CCC ACA GGG GAT	720
Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro Thr Gly Asp	
760 765 770	
GGC CAG AGC CTG GAA GAG CTC TCC TCC TGC TTC TAC ACT GTC ATC CCA	768
Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr Val Ile Pro	
775 780 785	
CAC AAC TTC GGC CGC AGC CGA CCC CCG CCC ATC AAC TCC CCT GAT GTG	816
His Asn Phe Gly Arg Ser Arg Pro Pro Pro Ile Asn Ser Pro Asp Val	
790 795 800 805	
CTT CAG GCC AAG AAG GAC ATG CTG CTG GTG CTA GCG GAC ATC GAG TTG	864
Leu Gln Ala Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu	
810 815 820	
GCG CAG ACC TTG CAG GCA GCC CCT GGG GAG GAG GAG GAG AAA GTG GAA	912
Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu Glu Glu Lys Val Glu	
825 830 835	
GAG GTG CCA CAC CCA CTG GAT CGA GAC TAC CAG CTC CTC AGG TGC CAG	960
Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln Leu Arg Cys Gln	
840 845 850	
CTT CAA CTG CTG GAC TCC GGG GAG TCC GAG TAC AAG GCA ATA CAG ACC	1008
Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr Lys Ala Ile Gln Thr	
855 860 865	
TAC CTG AAA CAG ACT GGC AAC AGC TAC AGG TGC CCA AAC CTG CGG CAT	1056
Tyr Leu Lys Gln Thr Gly Asn Ser Tyr Arg Cys Pro Asn Leu Arg His	

870	875	880	885	
GTT TGG AAA GTG AAC CGA GAA GGG GAG GGA GAC AGG TTC CAG GCC CAC				1104
Val Trp Lys Val Asn Arg Glu Gly Glu Gly Asp Arg Phe Gln Ala His				
	890	895	900	
TCC AAA CTG GGC AAT CGG AGG CTG CTG TGG CAC GGC ACC AAT GTG GCC				1152
Ser Lys Leu Gly Asn Arg Arg Leu Leu Trp His Gly Thr Asn Val Ala				
	905	910	915	
GTG GTG GCT GCC ATC CTC ACC AGT GGG CTC CGA ATC ATG CCA CAC TCG				1200
Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser				
	920	925	930	
GGT GGT CGT GTT GGC AAG GGT ATT TAT TTT GCC TCT GAG AAC AGC AAG				1248
Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys				
	935	940	945	
TCA GCT GGC TAT GTT ACC ACC ATG CAC TGT GGG GGC CAC CAG GTG GGC				1296
Ser Ala Gly Tyr Val Thr Thr Met His Cys Gly Gly His Gln Val Gly				
	950	955	960	965
TAC ATG TTC CTG GGC GAG GTG GCC CTC GGC AAA GAG CAC CAC ATC ACC				1344
Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Lys Glu His His Ile Thr				
	970	975	980	
ATC GAT GAC CCC AGC TTG AAG AGT CCA CCC CCT GGC TTT GAC AGC GTC				1392
Ile Asp Asp Pro Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val				
	985	990	995	
ATC GCC CGA GGC CAA ACC GAG CCG GAT CCC GCC CAG GAC ATT GAA CTT				1440
Ile Ala Arg Gly Gln Thr Glu Pro Asp Pro Ala Gln Asp Ile Glu Leu				
	1000	1005	1010	
GAA CTG GAT GGG CAG CCG GTG GTG GTG CCC CAA GGC CCG CCT GTG CAG				1488
Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln Gly Pro Pro Val Gln				
	1015	1020	1025	
TGC CCG TCA TTC AAA AGC TCC AGC TTC AGC CAG AGT GAA TAC CTC ATA				1536
Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln Ser Glu Tyr Leu Ile				
	1030	1035	1040	1045
TAC AAG GAG AGC CAG TGT CGC CTG CGC TAC CTG CTG GAG ATT CAC CTC				1584
Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu Leu Glu Ile His Leu				
	1050	1055	1060	
TAA				1587

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 528 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys

1	5	10	15
Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu	20	25	30
Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro	35	40	45
Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr	50	55	60
Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe	65	70	75
Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn	85	90	95
Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe	100	105	110
Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu	115	120	125
Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro	130	135	140
Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu	145	150	155
Ala Val Val Lys Val Asp Ser Gly Pro Val Arg Thr Val Val Lys Pro	165	170	175
Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn Ile Phe Ser	180	185	190
Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu Asp Val Lys	195	200	205
Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala Arg Gly Phe	210	215	220
Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro Thr Gly Asp	225	230	235
Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr Val Ile Pro	245	250	255
His Asn Phe Gly Arg Ser Arg Pro Pro Pro Ile Asn Ser Pro Asp Val	260	265	270
Leu Gln Ala Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu	275	280	285
Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu Glu Glu Lys Val Glu	290	295	300
Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln Leu Leu Arg Cys Gln	305	310	315
Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr Lys Ala Ile Gln Thr			



(2) INFORMATION FOR SEQ ID NO: 11:

(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

(A) NAME/KEY: Region  
(B) LOCATION:2  
(D) OTHER INFORMATION:/note= "Xaa steht fuer 1 bis 5  
andere Aminosaeuuren"

- 72 -

- (A) NAME/KEY: Region
- (B) LOCATION:3
- (D) OTHER INFORMATION:/note= "Xaa steht fuer Ser oder Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Pro Xaa Xaa Gly Xaa Xaa Xaa Gly Lys Gly Ile Tyr Phe Ala  
 1 5 10

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:1
- (D) OTHER INFORMATION:/note= "Xaa steht fuer Ser oder Thr"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:6
- (D) OTHER INFORMATION:/note= "Xaa steht fuer Ile oder Val"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:9
- (D) OTHER INFORMATION:/note= "Xaa steht fuer 1 bis 5 andere Aminosaeuren"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:10
- (D) OTHER INFORMATION:/note= "Xaa steht fuer Ser oder Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Xaa Xaa Gly Leu Arg Xaa Xaa Pro Xaa Xaa Gly Xaa Xaa Xaa Gly Lys  
 1 5 10 15  
 Gly Ile Tyr Phe Ala  
 20

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:

- 74 -

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: Region

(B) LOCATION:1

(D) OTHER INFORMATION:/note= "Xaa steht fuer Leu oder Val"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa
1          5          10          15
Xaa Xaa Xaa Xaa Xaa Leu
                20

```

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: Region

(B) LOCATION:21

(D) OTHER INFORMATION:/note= "Xaa steht fuer Asp oder Glu"

(ix) FEATURE:

(A) NAME/KEY: Region

(B) LOCATION:22

(D) OTHER INFORMATION:/note= "Xaa steht fuer 10 oder 11 andere Aminosaeuren"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

```

Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn Xaa Xaa Tyr Xaa Xaa
1          5          10          15
Gln Leu Leu Xaa Xaa Xaa Trp Gly Arg Val Gly
                20          25

```

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Ala	Xaa	Xaa	Xaa	Phe	Xaa	Lys	Xaa	Xaa	Xaa	Xaa	Lys	Thr	Xaa	Asn	Xaa
1				5				10						15	
Trp	Xaa	Xaa	Xaa	Xaa	Xaa	Phe	Xaa	Xaa	Xaa	Pro	Xaa	Lys			
				20				25							

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 44 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: Region  
(B) LOCATION: 4  
(D) OTHER INFORMATION: /note= "Xaa steht fuer Ile oder  
Leu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Gln	Xaa	Leu	Xaa	Xaa	Xaa	Ile	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5				10						15	
Met	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Pro	Leu	Gly	Lys	Leu
				20				25					30		
Xaa	Xaa	Xaa	Gln	Ile	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Leu
			35					40							

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Phe	Tyr	Thr	Xaa	Ile	Pro	His	Xaa	Phe	Gly	Xaa	Xaa	Xaa	Pro	Pro
1				5					10					15

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Lys	Xaa	Xaa	Xaa	Leu	Xaa	Xaa	Leu	Xaa	Asp	Ile	Glu	Xaa	Ala	Xaa	Xaa
1				5				10						15	

Leu

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Gly	Xaa	Xaa	Xaa	Leu	Xaa	Glu	Val	Ala	Leu	Gly
1				5					10	

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:14
- (D) OTHER INFORMATION:/note= "Xaa steht fuer 7 bis 9  
andere Aminosaeuren"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Gly Xaa Xaa Ser Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Pro Xaa Leu Xaa  
1 5 10 15  
Gly Xaa Xaa Val  
20

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:2
- (D) OTHER INFORMATION:/note= "Xaa steht fuer Tyr oder  
Phe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Glu Xaa Xaa Xaa Tyr Xaa Xaa Xaa Gln Xaa Xaa Xaa Xaa Tyr Leu Leu  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg Ala  
1 5 10 15  
Leu Asn Glu Ser

20

## (2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Lys	Thr	Glu	Leu	Gln	Ser	Pro	Glu	His	Pro	Leu	Asp	Gln	His	Tyr	Arg
1				5					10					15	
Asn Leu His Cys															
20															

## (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Cys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu	Glu	Asp	Pro	Phe	Arg	Ser	Thr
1				5					10					15	
Ala Glu Ala Leu Lys															
20															

## (2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:



